Validating Revised Bloom's Taxonomy Using Deep Knowledge Tracing

Amar Lalwani(✉) and Sweety Agrawal

funtoot, Bangalore, India
{amar.lalwani,sweety.agrawal}@funtoot.com

Abstract. Revised Bloom's Taxonomy is used for classifying educational objectives. The said taxonomy describes a hierarchical ordering of cognitive skills from simple to complex. The Revised Taxonomy relaxed the strict cumulative hierarchical assumptions of the Original Taxonomy allowing overlaps. We use a knowledge tracing model, Deep Knowledge Tracing (DKT), to investigate the hierarchical nature of the Revised Taxonomy and also study the overlapping behavior of the Taxonomy. The DKT model is trained on about 42 million problems attempted on funtoot by the students. funtoot is an adaptive learning platform where students learn by answering problems. We propose a novel way to interpret the model's output to measure the effects of each learning objective on every other learning objectives. The results confirm the relaxed hierarchy of the skills from simple to complex. Moreover, the results also suggest overlaps even among the non-adjacent skills.

Keywords: Deep knowledge tracing · Revised Bloom's Taxonomy
Cognitive skills · Hierarchical taxonomy · Deep learning
Student modeling · Domain knowledge · funtoot

1 Introduction

Benjamin S. Bloom, along with a group of educators took up a task of classifying educational goals and objectives. They aimed to classify thinking behaviors that were believed to be important in the process of learning. The output of their research was a taxonomy of three domains:

- The cognitive
- The affective
- The psychomotor

The cognitive domain was further broken down into six cognitive levels of complexity (called, learning objectives): Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation.

C. Penstein Rosé et al. (Eds.): AIED 2018, LNAI 10947, pp. 225–238, 2018.
https://doi.org/10.1007/978-3-319-93843-1_17

The levels are often depicted as a stairway, with emphasis on climbing to a higher level of complexity. The taxonomy is hierarchical, each level is subsumed by the higher levels. In other words, a student functioning at the 'Application' level has also mastered the material at the 'Knowledge' and 'Comprehension' levels.

A revision of this framework was developed in a similar manner 45 years later [2]. The Revised Taxonomy is also hierarchical in nature, like the original one. In the Revised Taxonomy, the six major categories - Remember, Understand, Apply, Analyse, Evaluate, and Create, differ in their complexity, with Remember being less complex than Understand, which is less complex than Apply, and so on. However, the complexity of the six categories is allowed to overlap, unlike in the case of Original Bloom's Taxonomy. Great emphasis is placed on the teacher usage rather than on developing a strict hierarchy.

Bloom's Taxonomy has been widely used in curriculum development [7], student assessment [23] and instruction evaluation. In curriculum design it serves as a common vocabulary to structure the curriculum's learning objective within the competency-based curriculum [3]. There are instructors who have applied Bloom's Taxonomy successfully in their classrooms. In circumstances when an instructor desires to move a group of students through a learning process utilizing an organized framework, Bloom's Taxonomy has been proven to be helpful. Research has shown that the implementation of Bloom's taxonomy in the curriculum has shown a positive outcome on students' test score [3,7,18].

Two important points, which various researchers have attempted to validate, have been noted about the original taxonomy, as described in [4,10]:

1. Taxonomy contained the categories ordered from simple to complex and from concrete to abstract.
2. The taxonomy was assumed to have a cumulative hierarchy; that is, mastery of a simple category was a prerequisite to the mastery of the next more complex category.

The authors have mentioned that even though the assumption of strict hierarchy is relaxed in the new taxonomy [2, Appendix A], they do not mean that the Revised Taxonomy is not hierarchical. Although the definition has been changed slightly, the authors believe that the empirical evidence found for the original taxonomy is not invalidated for the Revised Taxonomy. Hence, the empirical evidence found in [2, Appendix A] should apply to the Revised Bloom's Taxonomy as well.

Funoot [1] is a personalized learning tutor which employs Revised Bloom's Taxonomy to organize the domain knowledge. Each problem in funtoot is designed to cater to a specific learning objective as defined in the Revised Taxonomy. We model the students' interactions generated on funtoot using a deep knowledge tracing (DKT) model [19].

[1] http://www.funtoot.com/.

In this study, we use the DKT model to study the interconnections between different skills of the Revised Taxonomy. We then investigate the validity of the above two mentioned assumptions and explore how well those assumptions are realized from the student interaction data. The knowledge tracing model is modeled with *learning objectives* as features, which gives us immense flexibility to simulate a controlled environment. The different behaviors of the student like mastery and learning can be encoded and the effects of learning objectives in the process of learning can be investigated in detail.

The rest of the paper is organized as below: Sect. 2 discusses some of the important prior works. Section 3 describes the dataset used to train the knowledge tracing model. Section 4 describes the knowledge tracing model. Section 5 describes the experiments done. Section 6 reports the results of the experiment. Section 7 concludes the paper with future work.

2 Literature Survey

This section discusses some of the works which also attempt to verify the assumptions and/or claims of Bloom's Taxonomy as discussed in Sect. 1. Multiple attempts have been made in the past using various techniques of analysis to verify the hierarchy of Bloom's Taxonomy as well as Revised Bloom's Taxonomy.

A meta analysis of [5,6,9,11,17,22] was conducted by the authors of Revision of Bloom's Taxonomy [2, Chap. 16] to verify the cumulative hierarchical nature in the original Bloom's Taxonomy [4]. These studies were considered for the meta-analysis as they had published their original inter-correlation data, whereas, other similar studies that had not published such data were not included in this study.

The inter-correlation data consisted of the category scores for a sample of students who had taken the same tests. Each test measured a single taxonomic level of the original Bloom's Taxonomy. Each such score was then paired with the score of every other category and correlations of such pairs were computed.

In case if the cumulative hierarchy were present, success in one category would likely be accompanied by success in the category closest to it in the hierarchy. Success in one category is a necessary but not a sufficient condition for success in the more complex adjacent category. The correlations between these categories would be higher than with the still more complex categories. Also, successively more complex categories should show successively lower correlations [2, Chap. 16].

The study concluded that the hierarchical ordering holds for the middle categories - Comprehension, Application and Analysis. Authors had also reversed the last two categories of original taxonomy, Create and Evaluate to the revised taxonomy's, Evaluate and Create to see how the data would look like. The authors concluded that doing so gave somewhat better results.

The conclusion of the meta analysis was similar to the conclusion of most of the six studies individually, namely, that, excluding Knowledge, the support for a cumulative hierarchy was seen in the simpler categories.

Two other studies in [8,15] were done using the data from Kropp and Stoker's [11,20] work. Authors of [15] used step-wise regression, path analysis and factor analysis to reanalyse the data from Kropp and Stoker's [11,20] ninth grade sample of one of their tests - Atomic Structure. They found that all the techniques had rejected a simple hierarchical assumption of the taxonomy. Their results do not align with the analysis of the Kropp and Stoker's study. Authors of [8] applied an alternative method to the Kropp and Stoker's [11,20] data. They concluded that the simplex assumption of the Bloom's Taxonomy is supported when *Knowledge* is removed from the taxonomy.

Study in [12] showed that students are given a task that they can complete, and when memory is tested for the same material, students who operate at higher taxonomic level will produce superior memory scores than the students who operated on the lower taxonomic level. Their results conform and provide a moderately strong support for the hierarchical nature of the taxonomy.

A study done in [16] showed that the activities and assignments for Internet for Business class in the order of cognitive skills of the Revised Bloom's Taxonomy prepare students with higher order thinking skills. They showed that practicing the skills in the order of their difficulties had a positive effect on the higher order skills.

All the above mentioned works except for [15] showed some evidence in support for the hierarchical nature of the taxonomy. The meta-analysis [2, Chap. 16] and [8] both showed the support for the hierarchical nature after removing Knowledge level. Authors of [2] have also discussed the ambiguity of the placement of the Knowledge in the hierarchy based on the meta-analysis.

Madaus et al. have used squared semi-partial correlations and fitted a path model to the levels of the taxonomy in [14]. They found that not only did the magnitude of the paths between the adjacent levels declined as the levels became more complex in the Bloom's Hierarchy, but also that there were significant paths between non-adjacent levels. This results are similar to our findings for the Revised Bloom's Taxonomy presented in this paper.

Authors of [18] attempted to categorize the cognitive skills involved in interpreting medical images. The authors had hypothesized that hierarchical levels of cognitive process would emerge to define different levels of learning medical imaging concepts. The authors have used Revised Bloom's Taxonomy to assign a category to each of the questions of their examination mentioned in the study. The authors concluded that there was an inverse relationship between the depth of cognitive process and mean scores. With the increase in the complexity of the skills, there was a decrease in the score. However, the Revised Taxonomy's original script [2] says that it is not necessary to sequentially achieve each of the levels, that is, students can achieve Create before the Remembering process. This is consistent with the concept of problem-based learning [21], which emphasizes on the early immersion in diagnostic thinking [18].

Athanassiou et al. describes the usage of Bloom's Taxonomy in Management classroom in order to facilitate students with *Scaffolding* in [3]. Usage of Bloom's taxonomy in this fashion allows students to determine the level of her own work. The study concluded that the use of the taxonomy as a scaffolding device helped students improve their skills and become more aware of their learning process and hence, have a pointed way of making improvements. Study in [3,16] shows that the usage of Bloom's Taxonomy helps improve the higher order skills in students.

3 Dataset

In funtoot [13], the curriculum is designed such that a subject is broken down into multiple high level topics. These topics are then divided into smaller units called sub-topics (*sc*). *sc* is further broken down into a smallest teachable unit called sub-sub-topic (*ssc*). This *ssc* can not be broken down further and is the most granular level skill. For instance, subject *Mathematics* has a topic called *Addition*. Topic *Addition* is further broken down into *scs* - *Addition of two digit numbers*, *Addition of three digit numbers* and so on. The *Addition of two digit numbers* contains smallest teachable units - *sscs*, like *Addition of two digit numbers with carry over*, *Addition of two digit numbers without carry over* and so on.

sscs are designed to achieve an expert defined learning objective. Problems are then designed to cater to these learning objectives. A Bloom's Taxonomy learning objective (*btlo*) tag is assigned to each problem based on the learning objective (from the Revised Bloom's Taxonomy) it helps to achieve. We call this a *btlo* tagging. *ssc* may not always have all the six learning objectives as defined in the Revised Bloom's Taxonomy.

Funtoot is a classroom based personalized learning tutor for grades 2 to 9 for subjects - Maths and Science. It is used in over 100 schools following one of the boards of education[2] like: CBSE, KASB, ICSE or IGCSE. There are broadly 74 topics spanning Maths and Science in funtoot. These 74 topics have a total of 682 *sscs*.

In this study, we wish to verify the validity of the claims of the Revised Bloom's Taxonomy and Bloom's Taxonomy. Hence, it is important for us to have at least two *btlos* in an *ssc* to check the effect of one *btlo* on another. For this reason, we have only considered *sscs* with more than two *btlos* which leaves us with 536 *sscs*. Figure 1 shows the distribution of a number of *sscs* for *btlos*.

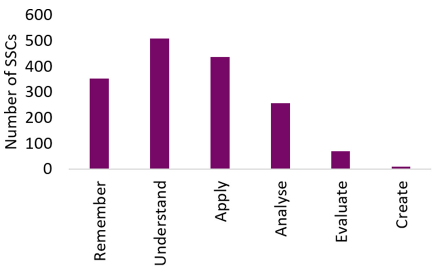

Fig. 1. *ssc* distribution across *btlos*

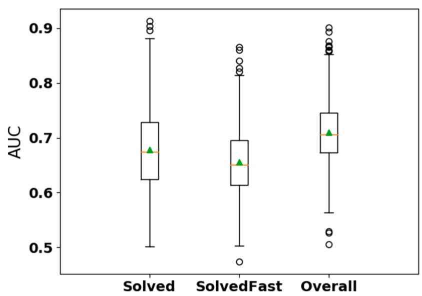

Fig. 2. AUCs

[2] https://en.wikipedia.org/wiki/Boards_of_Education_in_India.

The dataset has a total of 41.7 million problem interaction entries generated by 1,03,593 students while interacting with 10,158 problems. Each interaction represents a problem attempted by the student which is characterized by two attributes: *solved-fast* (got the correct answer in first attempt) and solved (got the correct answer in one or more attempts). While attempting the problem, if the student's response is incorrect, the student is notified about it along with the explanation about why it is incorrect. She then gets another attempt at the problem with a hint (if available). This continues till the predefined number of attempts for that particular problem are exhausted. In the end, a student is also provided with the detailed solution of the problem. Out of 41.7 million data points, 29.7 million are solved (71%) and 22.7 million are solved-fast (54%).

4 Deep Knowledge Tracing (DKT)

Intelligent Tutoring systems like funtoot aim to provide personalized study material to students and help when needed. For this, the knowledge state of the student needs to be monitored frequently and updated as and when students interact with a digital tutor. The models capable of doing this are called Knowledge tracing models. By modeling knowledge acquisition process of the students, predictions can be made on the future interactions and based on these predictions further instructions can be delivered in a personalized manner.

Several knowledge tracing models exist in literature. For this study, we have used a Deep knowledge Tracing (DKT) model proposed in [19]. DKT is a time series model based on Recurrent Neural Network. It takes as input the series of exercises and a bit corresponding to it indicating the outcome from the student interactions with the tutor.

In this study, we want to measure the effect of one *btlo* on its successor in the Revised Bloom's Taxonomy. We need a scope within which the DKT model can capture interlinks between the *btlos*. In funtoot's knowledge hierarchy, the *btlo* is present in the *ssc* scope. Hence, we have trained the models at an *ssc* level for simplicity.

Since we have 536 *sscs*, we have 536 DKT models and *btlos* are used as features. So, the input vector is a series of *btlos* (tagged to the problems attempted by the student) and their corresponding outcomes of the problem attempts: solved and solved-fast in our case. For instance, if an *ssc* consists of three *btlos*, the input consists of three neurons per *btlo*, one representing whether the interaction belonged to that *btlo*, one for solved and one for solved-fast. The model outputs the probability that the problem will be solved and solved-fast for each *btlo*, amounting to two output neurons per *btlo*.

The average AUC of the DKT models is 0.69 ($\sigma = 0.07$) for *solved*, 0.66 ($\sigma = 0.05$) for *solved-fast* and 0.71 ($\sigma = 0.06$) overall. Figure 2 shows the boxplot of the distribution of AUCs across *sscs* for outputs solved, solved-fast, and overall (average of solved and solved-fast).

5 Experiments

The aim of this study is to verify if the skills in the Revised Bloom's Taxonomy are ordered from lower to higher level. We also attempt to see if lower order skills are the prerequisites for higher order skills. In other words, we want to study the effects of mastery of lower order skills on the chances of mastery of higher order skills. We hypothesize that student's *mastery* in the lower order skills should increase her chances of mastery in the higher order skills. The evidence of *mastery* of any skill is shown by solving successively 5 problems from that skill in the first attempt (solved-fast).

We measure the effect of a lower order skill i on the higher order skill j in following ways. Formula 1 measures the change in the solved-fast probability of skill j after having shown the evidence of mastery of skill i compared to the initial state (*).

$$(i \rightarrow j) - (* \rightarrow j) \tag{1}$$

Here, * is the initial state when none of the skills are mastered and none of the problems are solved by the students. This serves as the starting state of the student.

Formula 2 shows the combined effect of skill i and all its lower skills on skill j.

$$(\widehat{i} \rightarrow j) - (* \rightarrow j) \tag{2}$$

Here, $\widehat{i}$ is sequence of skills easier than i, including i, in simple to complex order. In this formula, skill $\widehat{i}$ is mastered and the change of its effect on the probability of skill j is measured compared to the initial state *. The mastery of such a set of skills $(\widehat{i})$ is shown by showing the mastery of each skill in the same order.

Formula 1 represents the effect of skill i on skill j, which might include the effects of the skills easier than i, whereas Formula 2 measures the combined effect of skills $\widehat{i}$ on skill j. Both the above mentioned formulae do not show the individual effect of skill i after the mastery of the lower order skills is achieved. We compute this by Formula 3.

$$(\widehat{i} \rightarrow j) - (\overline{i} \rightarrow j) \tag{3}$$

Here, $\overline{i}$ is the set of the skills lower to skill i in the simple to complex order. Formula 3 shows the isolated effect of skill i on skill j by taking the difference of probability of skill j when skills $\widehat{i}$ and $\overline{i}$ are mastered.

For instance, consider an *ssc* containing four *btlos*: *Remember, Understand, Apply* and *Analyse*. When we say $\widehat{Apply}$, we mean the sequence of skills *Remember, Understand, Apply*. When we say $\overline{Apply}$, we mean the sequence of skills *Remember, Understand*.

If a lower order skill i is a prerequisite of a higher order skill j, then the mastery of the skill j would assure the mastery of skill i. We study this prerequisite nature by measuring the effect of the mastery of the skill j on the mastery of skill i through the following formulas.

In Formula 4, the probability of the skill i after the mastery of skill j is compared against the probability of skill i at the initial state.

$$(j \longrightarrow i) - (* \longrightarrow i) \tag{4}$$

If indeed skill i is the prerequisite of skill j, the probabilities of skill i after the mastery of skill j would be the same as the probability of skill i after having mastered the skill i. The difference between these probabilities is captured by Formula 5.

$$(j \longrightarrow i) - (i \longrightarrow i) \tag{5}$$

The skills lower than i might also have an effect on the mastery of skill i which is not considered by the Formula 5. Hence, in Formula 6, the probability of skill i after the mastery of skill j is compared against the probability of skill i after the mastery of the skills $\widehat{i}$.

$$(j \longrightarrow i) - (\widehat{i} \longrightarrow i) \tag{6}$$

We also compare the probabilities of each pair of skills at the initial state and the final state (when all the skills are mastered in simple to complex order). For this, we have measured the correlations between the probabilities of all pairs of skills and the differences between them are also computed.

The results of the initial and final correlations are shown in Tables 2 and 3 respectively. Tables 4 and 5 show the difference between the probabilities at the initial and final state. The results of all the above formulae 1, 2, 3, 4, 5 and 6 are shown in Tables 6, 7, 8, 9, 10 and 11 respectively.

In Tables 6, 7 and 8, we measure the effect $(i \rightarrow j)$. Here i is the row header and j is the column. That is, the top left most cell $(1, 1)$ refers to $(R \rightarrow U)$ in these tables. In Tables 9, 10 and 11, we measure the effect $(j \rightarrow i)$. Here j is the row header and i is the column. That is, the top left most cell $(1, 1)$ refers to $(U \rightarrow R)$ in these tables.

To compare a pair of skills, the *sscs* having both the skills are used and the number of *sscs* for all the pairs of skills is shown in Table 1. In formula 2, 3 and 6, we consider the *sscs* where at least one lower skill than i is present. The reduced number of *sscs*, if any, for all such pairs are shown in the parentheses in Table 1.

Table 1. Counts of SSCs containing BTLO Pairs

	U	Ap	An	E	C
R	339	264	165	39	3
U		411(253)	237(160)	64(39)	7(3)
Ap			222(209)	63(61)	8(7)
An				64(62)	8
E					7

Table 2. Initial correlations

	U	Ap	An	E	C
R	0.39(0.0)	0.14(0.03)	0.1(0.22)	−0.24(0.15)	−0.58(0.61)
U		0.34(0.0)	0.28(0.0)	0.16(0.21)	0.83(0.02)
Ap			0.22(0.0)	−0.07(0.6)	−0.23(0.58)
An				0.36(0.0)	0.07(0.86)
E					0.68(0.09)

6 Results

In Tables 2 and 3, the values in the parentheses are the *p-values*. From Table 2, it can be seen that for the skills *Remember*, *Understand*, *Apply* and *Analyse*, their initial correlations with their adjacent more complex skills is the highest and the correlations decrease with the increase in the difficulty of the skills. The correlations at the diagonal cells decreases with the increase in the complexity of the skill. However, there is an increase in the correlation of the pair *Analyse-Evaluate* compared to the previous diagonal pair *Apply-Analyse*. Even though the correlation between the pair *Evaluate-Create* is very high, its *p-value* is 0.09. Ideally, for all the columns, the maximum correlations should be the diagonal entries. This pattern holds for all the *btlos* except for *Analyse* and *Create*. Both *Analyse* and *Create* achieve maximum correlation of 0.28 and 0.83 respectively, with the same *btlo Understand*. The initial differences are negative for the pairs where complex skills are involved in Table 4. The column entries for *Create* are the most negative values.

The pattern where each entry in the matrix is less than the entry to its left and also to the entry below it, is called *simplex* [2, Chap. 16]. The adherence to this simplex behavior in initial correlations is quite prominent. However, this is not the case with final correlations as seen in Table 3. The correlations of skills *Understand* and *Apply* with the complex skills increases with the increase in the complexity of the skills. The first three diagonal correlations see a decrease with the increase in the complexity of the skills. The correlation between the pair *Evaluate* and *Create* is 0.74 but it is marginally significant (*p-value*= 0.06). Like the initial differences, the final differences (Table 5) between the probabilities are negative for the complex skills.

In the rows of the skills *Remember*, *Understand* and *Apply* in Table 6 of Formula 1, the values decrease with the increase in the complexity of the skills except for the last column *Create*. The effect of the skills *Understand*, *Apply* and

Table 3. Final correlations

	U	Ap	An	E	C
R	0.2 (0.0)	0.08 (0.19)	0.04 (0.64)	−0.21 (0.2)	−1.0 (0.02)
U		0.18 (0.0)	0.11 (0.08)	0.27 (0.03)	0.78 (0.04)
Ap			0.14 (0.04)	0.25 (0.05)	0.77 (0.03)
An				0.11 (0.39)	0.49 (0.22)
E					0.74 (0.06)

Table 4. Initial differences

	U	Ap	An	E	C
R	0.02	0.03	0.01	0.03	−0.02
U		0.01	−0.01	−0.02	−0.11
Ap			−0.03	−0.04	−0.09
An				−0.02	−0.05
E					−0.01

Table 5. Final differences

	U	Ap	An	E	C
R	0.07	0.1	0.04	0.12	0.05
U		0.03	−0.01	−0.04	−0.16
Ap			−0.06	−0.07	−0.07
An				−0.05	−0.03
E					0.01

Table 6. Formula 1

	U	Ap	An	E	C
R	0.1	0.04	0.03	0.01	−0.08
U		0.05	0.03	0.02	0.06
Ap			0.07	0.03	0.08
An				0.05	0.1
E					0.14

Analyse on the skill *Create* is even greater than their effects on their respective adjacent complex skills, *Apply*, *Analyse* and *Evaluate*. The skill *Evaluate* has the maximum effect of 0.14 on its adjacent complex skill, *Create*.

The effect of $\widehat{Understand}$ and $\widehat{Apply}$ is maximum on their respective complex adjacent skills, *Apply* and *Analyse* computed using Formula 2 (seen in Table 7). However, the effect of $\widehat{Apply}$ and $\widehat{Analyse}$ on the skill *Create* is greater than or comparable to their effects on their respective complex adjacent skills - *Analyse* and *Evaluate*. The effect of $\widehat{Evaluate}$ on its complex adjacent skill *Create* is the maximum effect in the table.

The isolated effects (computed using Formula 3) of the skills in Table 8 are similar to the effects seen in the Table 7 with a slight reduction in the magnitude. However, there a couple of exceptions where we see a significant change. The isolated effect of *Understand* on the skill *Create* is the maximum (0.1) and the isolated effect of the skill *Analyse* is comparatively lesser on skill *Create*.

Formula 4 captures the effect of each skill on its lower order skills, the results of which can be seen in Table 9. Ideally, the diagonal entries should decrease with the increase in the complexity of the skill. Moreover, the entries in each row should increase as they move further away from the diagonal. Also, the column entries should increase as they move away from the diagonal. This expected behavior is observed very clearly with a few exceptions. The effects of *Create* on *Understand* and *Apply* are lower than expected and the effect of *Create* on *Evaluate* is much higher than expected.

Table 7. Formula 2

	U	Ap	An	E	C
R	-	-	-	-	-
U		0.07	0.07	0.02	0.02
Ap			0.09	0.04	0.08
An				0.07	0.09
E					0.1

Table 8. Formula 3

	U	Ap	An	E	C
R	-	-	-	-	-
U		0.03	0.04	0.01	0.1
Ap			0.04	0.02	0.06
An				0.02	0.02
E					0.07

Table 9. Formula 4

	R	U	Ap	An	E
U	0.12				
Ap	0.13	0.11			
An	0.14	0.11	0.09		
E	0.15	0.12	0.09	0.06	
C	0.18	0.1	0.11	0.14	0.13

Table 10. Formula 5

	R	U	Ap	An	E
U	−0.08				
Ap	−0.06	−0.05			
An	−0.04	−0.03	−0.03		
E	−0.03	−0.05	−0.03	−0.07	
C	0.12	−0.08	0.04	−0.02	0.03

If the effect of a skill j on a simpler skill i computed using Formula 5 happens to be close to zero, that would imply that the skill j completely subsumes skill i. For the cumulative hierarchical assumptions to hold true for the Revised Bloom's Taxonomy, all the values in Table 10 should have been close to zero, which is not the case here.

Table 11. Formula 6

	R	U	Ap	An	E
U	-				
Ap	-	−0.07			
An	-	−0.06	−0.05		
E	-	−0.03	−0.07	−0.1	
C	-	−0.04	0.02	−0.04	0.06

The effects computed using Formula 6, shown in Table 11 are similar to the effect seen in the earlier Table 10 with little reduction in the magnitude. However, the effects of *Evaluate* on *Understand* and the effects of *Create* on *Understand* and *Evaluate* show an unexpected increase. Ideally, all the effects in the Table 11 should have been similar to the effects in the Table 10 if the higher skills were to subsume the lower skills.

7 Discussion and Conclusion

Based on the correlations and the results of Formulae 1, 2 and 3 (Tables 1, 2 and 3), it can be clearly said that skill *Understand* is more complex than skill *Remember* and skill *Apply* is more complex than skill *Understand*. Skill *Apply* has good effect on skill *Analyse* and their probabilities share a moderate correlation indicating the complexity of *Analyse* being more than *Apply*. However, due to the higher complexity of the skill *Analyse*, this effect of *Apply* on *Analyse* is not expected to be very high.

There is a moderate initial correlation of *Analyse* with *Evaluate*. The effect of *Analyse* on *Create* is higher than expected. The effect of *Evaluate* on *Create* is exceptionally good. These points strongly suggest that the difference in complexity between the higher order skills *Analyse*, *Evaluate* and *Create* is not very high.

The correlation of *Understand* with *Analyse* and high effect of *Understand* on *Analyse* indicate some overlap between the two skills. Both *Understand* and *Apply* have significant effects on *Create* which are even higher than their effects on their respective adjacent complex skills, *Apply* and *Analyse*. This hints some overlap between the skills *Understand*, *Apply* and *Create*. The overlap between *Understand* and *Apply* has also been admitted in the script of Revised Bloom's Taxonomy [2].

From the above evidence, *Understand* seems to correlate with almost all the higher order skills. *Understand* might be force fitted and seems out of place. It seems to not have any individual and unique place in the relative ordering of the skills. *Understand* was not included in the Original Taxonomy for the very similar reason that, *Understand* for all the practical purposes actually means anything from *Comprehension* to *Synthesis*. Despite of that, *Understand* was included in the Revised Taxonomy considering its widespread usage as a synonym of *Comprehension*.

Barring few exceptions in the results in Table 9, it can be clearly said that the skills are rightly ordered from simple to complex in the Revised Taxonomy. These results are based on the effects of mastery of skills on their lower order skills ($j \rightarrow i$) compared to the initial states. Hence, mastery of higher order skills always increased the likelihood of the mastery of simpler skills, with the larger likelihoods for more simpler skills. The exceptions suggested some overlap between the pairs *Create*-*Understand*, *Create*-*Apply* and *Create*-*Evaluate*.

However, when the effects of mastery of skills on its lower order skills ($j \rightarrow i$) are studied against the effects of already mastered lower order skills on themselves ($i \rightarrow i$), the results (as seen in Tables 10 and 11) deny the subsumption of the lower order skills by the higher order skills. The chances of the mastery of the lower order skills in the consequence of the mastery of higher order skills do increase, but not to the extent of considering lower order skills as the prerequisites (complete subset) of higher order skills.

In our attempt to generate optimal problem sequences using Revised Bloom's Taxonomy in [1], we found that when the problems are given to the students in the order of the hierarchy, the gains produced were higher when measured using Deep Knowledge Tracing model. These findings indicate that the Revised Bloom's Taxonomy is indeed hierarchical. Similar results were reported in [16] where students practiced the skills in the order of their difficulties (of the Revised Bloom's Taxonomy) and it had a positive effect on the higher order skills.

Our results do not align with the findings of the work [18] which showed that their implementation of Revised Bloom's Taxonomy had a cumulative hierarchy. We conclude that the Revised Bloom's Taxonomy is indeed a hierarchy (clearly observed in the lower learning objectives *Remember*, *Understand* and *Apply*) when the skills are judged on their relative (median) complexity. But, it is not a strict hierarchy in the sense that it allows for overlaps even among the non-adjacent skills, ensuring higher order skills do not subsume lower order skills. Mastery of higher learning objective does not assure mastery of lower learning objectives.

References

1. Agrawal, S., Lalwani, A.: Analysing problem sequencing strategies based on revised Bloom's taxonomy using deep knowledge tracing. In: 14th International Conference on Intelligent Tutoring Systems, June 2018. (to appear)
2. Anderson, L.W., Krathwohl, D.R., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., Wittrock, M.: A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy. Longman Publishing, New York (2001). Artz, A.F., Armour-Thomas, E. (1992). 9(2), 137–175
3. Athanassiou, N., McNett, J.M., Harvey, C.: Critical thinking in the management classroom: Bloom's taxonomy as a learning tool. J. Manage. Educ. **27**(5), 533–555 (2003)
4. Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R.: Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain, vol. 19. David McKay Co. Inc., New York (1956)
5. Davis, F.B.: Research in comprehension in reading. Reading Res. Q. **3**(4), 499–545 (1968)
6. Hancock, G.R.: Cognitive complexity and the comparability of multiple-choice and constructed-response test formats. J. Exp. Educ. **62**(2), 143–157 (1994)
7. Hawks, K.W.: The effects of implementing Bloom's taxonomy and utilizing the virginia standards of learning curriculum framework to develop mathematics lessons for elementary students (2010)
8. Hill, P., McGaw, B.: Testing the simplex assumption underlying Bloom's taxonomy. Am. Educ. Res. J. **18**(1), 93–101 (1981)
9. Klein, M.F.: Use of taxonomy of educational objectives (cognitive domain) in constructing tests for primary school pupils. J. Exp. Educ. **40**(3), 38–50 (1972)
10. Krathwohl, D.R.: A revision of Bloom's taxonomy: an overview. Theory Pract. **41**(4), 212–218 (2002)
11. Kropp, R.P., Stoker, H.W., Bashaw, W.: The validation of the taxonomy of educational objectives. J. Exp. Educ. **34**(3), 69–76 (1966)
12. Kunen, S., Cohen, R., Solman, R.: A levels-of-processing analysis of bloom's taxonomy. J. Educ. Psychol. **73**(2), 202 (1981)
13. Lalwani, A., Agrawal, S.: Few hundred parameters outperform few hundred thousand? Educational Data Mining (2017)
14. Madaus, G.F., Woods, E.M., Nuttall, R.L.: A causal model analysis of Bloom's taxonomy. Am. Educ. Res. J. **10**(4), 253–262 (1973)
15. Miller, W.G., Snowman, J.: Application of alternative statistical techniques to examine the hierarchical ordering in Bloom's taxonomy. Am. Educ. Res. J. **16**(3), 241–248 (1979)
16. Nkhoma, M.Z., Nkhoma, M.Z., Lam, T.K., Lam, T.K., Sriratanaviriyakul, N., Sriratanaviriyakul, N., Richardson, J., Richardson, J., Kam, B., Kam, B., et al.: Unpacking the revised Bloom's taxonomy: developing case-based learning activities. Education + Training **59**(3), 250–264 (2017)

17. Pachaury, A.: An empirical validation of taxonomy of educational objectives using McQuitty's hierarchical syndrome analysis. Indian Educ. Rev. **6**(2), 156–164 (1971)
18. Phillips, A.W., Smith, S.G., Straus, C.M.: Driving deeper learning by assessment: an adaptation of the revised Bloom's taxonomy for medical imaging in gross anatomy. Acad. Radiol. **20**(6), 784–789 (2013)
19. Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L.J., Sohl-Dickstein, J.: Deep knowledge tracing. In: Advances in Neural Information Processing Systems, pp. 505–513 (2015)
20. Stoker, H.W., Kropp, R.P.: Measurement of cognitive processes. J. Educ. Meas. **1**(1), 39–42 (1964)
21. Subramaniam, R.: Problem-based learning: concept, theories, effectiveness and application to radiology teaching. J. Med. Imaging Radiat. Oncol. **50**(4), 339–341 (2006)
22. Thomas, A.M.: Levels of Cognitive Behavior Measured in a Controlled Teaching Situation. Graduate School of Cornell University (1965)
23. Thompson, E., Luxton-Reilly, A., Whalley, J.L., Hu, M., Robbins, P.: Bloom's taxonomy for CS assessment. In: Proceedings of the Tenth Conference on Australasian Computing Education, vol. 78, pp. 155–161. Australian Computer Society, Inc. (2008)